Documentation to accompany the U.S. Geological Survey National Coal Resources Data System UStratigraphic (USTRAT) database

Introduction

Assessment of energy resources, particularly for the abundant coal resources within the U.S., is part of the scientific mission of both the U.S. Geological Survey (USGS) and State geological surveys. Traditional methods of coal resource calculation (Wood and others, 1983) require the compilation of point-source observations on the depth and thickness of the resource being studied. The UStratigraphic (USTRAT) database, part of the National Coal Resources Data System (NCRDS), preserves and provides access to publicly available stratigraphic descriptions from U.S. coal fields. Some point locations are contributed by USGS staff, but the majority is entered by federally-funded cooperating geologists from State agencies. These data have existed in digital format since 1975, but the purpose of this compilation is to facilitate the availability and use of these data by the public. This document provides background information on the data sources and attributes within USTRAT.

Database contents

Overview

The USTRAT database is unique in its breadth of coverage and public availability. Most of the coal-bearing States within the U.S. (Tully, 1996) are represented by information from both surface and subsurface coal stratigraphic descriptions. State stratigraphic nomenclature, which includes the use of informal names, is generally used. Many sources, including publications, exploration drilling, as well as field studies comprise the compilation. Some sources date back to the early twentieth or even late nineteenth century.

Data are stored in the same structural format of two tables having a one-to-many relationship. A header or point table has sample location and source information and a unit table contains multiple top-down descriptions of geologic units for that location. Some point locations have been entered by both USGS and State personnel, so duplication is possible, as well as multiple stratigraphic interpretations of the same point. See Table 1 for a list of attributes by table and the metadata file (http://weblink) for more detailed descriptions and sources of all attributes. The 'point id' field is the unique database identifier that links the header and unit tables.

The geographic distribution of points is not uniform because the sequence of data collection depended on availability and locally prioritized resource areas. The contents of USTRAT will be updated approximately on quarterly intervals each year, or as more data become publicly available.

Table 1a. Point attributes and explanation.

Point id	point id- unique number that links header and unit tables
Point name	point name / site name
Map name	map name (quadrangle)
Scale	scale of quadrangle map
State	state name
County	county name
Province	coal province name
Region	coal region name
Coal field	coal field name
Collecting Org	organization of collector
Collector	collector - name of submitter of point
Source	source of information
Surface elevation	surface elevation (feet)
Elevation precision	precision of surface (ground) elevation value
Total depth logged	total depth logged in feet
Point type	point type indicating type of data source
Number of units	number of stratigraphic units in this point
Prev collector	previous collector of same point
Prev point name	previous point name from previous collector or secondary name
Estimated rank	estimated rank of coal in point
Latitude	latitude in decimal degrees (N is positive)
Longitude	longitude in decimal degrees (W is negative)
Lat/long precision	latitude/longitude precision of location
Point date	associated date (year month day format - YYMMDD)
Principal meridian	principal meridian for township and range values
Quarter 1	quarter 1 (smallest)
Quarter 2	quarter 2
Quarter 3	quarter 3
Quarter 4	quarter 4 (largest)
Section	section within township and range
Township (N/S)	township number and north/south direction
Range (E/W)	range number and east/west direction
Weathering	weathering indicated, if present
Loc. strike	local strike
Loc. dip	local dip
Loc. angle	local angle
Project info	project information for groups of data submitted together
Point usage	point usage code
Comment	point comment containing related information
Ownership	ownership (F-federal, P-private, S-State, O-other) plus name
Hydro area	USGS hydrologic area code

Other database	other NCRDS databases that contain related data
Release date	date (if known) of public release
Last update	date information was last updated
Ownership	ownership (F-federal, P-private, S-State, O-other plus name)
Point notes	miscellaneous notes on the point (up to 250 characters)

Table 1 b- Unit attributes and explanation

Point id	point id- unique number that links header and unit tables
Unit number	unit number - 1 at top and numbered consecutively to bottom
Unit qualifier	unit qualifier letter code
From depth	from depth of unit (top) in feet
To depth	to depth of unit (bottom) in feet
Thickness	thickness of unit in feet
Formation	formation name (often informal usage and can contain group name)
Bed	bed name (often informal usage)
Lithology	lithology (see abbreviations list in appendix)
Lithology modifier	lithology modifier (see abbreviations list in appendix)
Color	color (see abbreviations list in appendix)
Grain size	grain size (see abbreviations list in appendix)
Grain shape	grain shape (see abbreviations list in appendix)
Mineral	mineralogy in unit (see abbreviations list in appendix)
Bedding	bedding
Contact	contact (see abbreviations list in appendix)
Fossil	Fossil presence in unit indicated by letter (i.e. P for plant)
Fracture	F indicates fractures in unit
Joint	J indicates joint presence in unit
Cleat	C indicates cleat in coal units (see comment field for unit)
Misc char	miscellaneous character used by point submitter
Misc num	miscellaneous number used by point submitter
Comment	unit comment - often tied to unit qualifier or cleat fields
Subzone	sub zone within coal zone
Resource zone	resource zone assigned for usage in resource calculations
Unit usage	unit usage indicates limitations on usefulness of point
Analytical id	analytical identifier in other NCRDS database
Last update	Date of last editorial revision to unit
Unit key	number relating to previous database identifier

Sources of Data

Every point in the database denotes its source – both as to the agency and person who submitted it, as well as the type of record on which it is based (e.g. core, outcrop, geophysical log). The public version of USTRAT currently includes data from: the USGS, the Geological Survey of Alabama, the Arkansas Geological Survey, the

Colorado Geological Survey, the Illinois State Geological Survey, the Indiana Geological Survey, the Iowa Geological Survey, the Kentucky Geological Survey, the Montana Bureau of Mines and Geology, the University of North Carolina (Chapel Hill), the University of North Dakota, the North Dakota Geological Survey, the New Mexico Bureau of Geology and Mineral Resources, the Ohio Division of Geological Survey, the Oklahoma Geological Survey, the University of Tennessee at Chattanooga, the Texas Bureau of Economic Geology, the West Virginia Geological and Economic Survey, and the Wyoming State Geological Survey.

Data quality and validation

The NCRDS database has had stratigraphic data entered by USGS scientists and state cooperators for over thirty years. The database has been migrated numerous times to different database backends utilizing data entry applications developed and modified over this time. Various methods of directly adding subsets of data to the database have also occurred over the past thirty years. The many different methods of data entry into the database and movement of the data into different database backends over time has introduced the possibility that errors have occurred in the transfer of data. Before releasing the data to the public, the NCRDS data will undergo data quality and validation checks to identify and eliminate potential errors within the database.

Additionally, different data sources will have inherently different errors in thickness and depth measurements (i.e. core versus cuttings versus geophysical log). All data have been subjected to validation procedures at both State and federal levels. In addition, some attributes have 'qualifiers' (e.g. surface elevation precision, latitude-longitude location precision, unit qualifier --- see http://XXXX) that will alert the user to issues on precision or data quality.

Before NCRDS data are released to the public, automated processes perform a number of checks on the data. The checks are divided into 2 categories: data attribution checks and spatial orientation checks.

The following data attribution checks are performed on the NCRDS data:

1. Each point is checked to see if it is has a corresponding unit record. Each point must have at least one corresponding unit record.

2. Each point must have a combination of point name, map name, and collector that uniquely identifies the point. Points that do not meet these criteria are deemed to be duplicate points.

3. The unit numbers associated with each point must start with 1 and be incremented by 1, so that a sequential order is established for the units associated with a point. Since not all units have depth information, the units' top and/or bottom depths can NOT be used to order them correctly.

4. The Point and Unit tables within the database are checked for broken referential constraints with values in the lookup tables associated with the database.

5. The value of the total depth logged (TDL) stored in the point table is verified by comparing the current value against the value derived from the summation of thickness of units associated to each point. The TDL should equal the sum of the associated unit thickness. A 1.5 percent difference between the sum of thicknesses for geologic units within a point and the stored header value for 'total depth logged' is considered an acceptable error.

6. The Point and Unit tables are checked for invalid data entries:

a. Invalid county – Does the county exist within state that has been entered?

b. Invalid Point Date – Is the date valid? The date is expected to be in "yyyymmdd" format.

c. Invalid proprietary code.- Is there an entry identifying the point as a proprietary point? Valid values are "PROP" for proprietary and Null or empty for non proprietary point.

d. Invalid quarters. – Is the Public Land Survey System quarter data entered in the correct order? Quarter 4 represents the largest quarter and must be entered first; therefore, quarter 1, being the smallest, entered last. For example, in order to store "NE of SW of NW" for quarters, "NW" should be entered for quarter 4, "SW" for quarter 3, "NE" for quarter 2, and nothing for quarter 1.

e. Units associated to a point – The number of units stored within the points table is verified by comparing with the actual count of associated units found in the unit table.

f. Point table is checked for invalid entries for fracture, joint, and cleat codes which are combined and stored in a single column in the point table. Possible codes are F for fracture, J for joint, C for cleat. The codes must be entered in the set order of first character = fracture, second character = joint, and third character = cleat. For example, if only a joint exists in a unit, entry of " J " (a blank space, J, and a blank space) is expected. Note that this field has been parsed to three separate columns in downloaded EXCEL files from the website.

Once the NCRDS data have passed the data attribution tests, a number of tests are run to determine whether the spatial information associated with each point is correct. The latitude and longitude coordinates for each point are analyzed for the following:

a. Are the point latitude-longitude coordinates within the US boundary?

b. Are the point latitude-longitude coordinates within the State boundary of the State name entered?

c. Are the point latitude-longitude coordinates within the county boundary of the county entered?

d. Are the point latitude-longitude coordinates within the topographic quadrangle boundary that matches the quadrangle name?

f. Are the point latitude-longitude coordinates within the township and range boundary of the township and range entered?

Upon passing the data attribution and spatial orientation checks, NCRDS public USTRAT data will be migrated to the NCRDS public web server where the data can be queried and downloaded through a mapping or tabular interface.

Database Access and Use

Searching for information

The first step is to decide whether to search for data graphically through a mapping query (http://energy.er.usgs.gov/ncrdstest/ncrds_flexapp/) or through a tabular query on attributes (http://energy.er.usgs.gov/ncrdstest/ncrds_data/). The mapping query allows you to select data by drawing a rectangle or polygon on the screen, with typical zoom and pan options. For a tutorial on this application, go to http://XXXX. A tabular attribute search allows the user to search and download data into spreadsheet format. A tutorial for this method is at (http://XXXX). Using the metadata file (http://XXXX) to understand attribute definitions will enable the user to increase the ease of data retrieval for either application.

Disclaimer

The data secured from this USGS Data Base are provisional and subject to revision. The data are released on the condition that neither the USGS nor the United States Government may be held liable for any damages resulting from its authorized or unauthorized use.

Distribution Liability

Although the data have been processed on the computer system at the U.S. Geological Survey, U.S. Department of the Interior, no warranty, expressed or implied, is made by the Geological Survey regarding the utility of the data on any other system, nor shall the act of distribution constitute any such warranty. No responsibility is assumed by the USGS in the use of these data.

Citing use of the database

Acknowledgments

References Cited

Wood, G.H., Jr., Kehn, T.M., Carter, M.D., and Culbertson, W.C., 1983, Coal resource classification system of the U.S. Geological Survey: U.S. Geological Survey Circular 891, 65 p.

Tully, John, 1996, Coal fields of the conterminous United States: U.S. Geological Survey Open-File Report 96-92, available online at http://pubs.er.usgs.gov/usgspubs/ofr/ofr9692 (Accessed, 1/10/2010)

Appendix of abbreviations

ABND	Abandoned
ABV	Above
ABDT	Abundant
ALUV	Alluvial
ALT	Altered(-ing)
AMT	Amount
ANAL	Analysis(-tical)
ANG	Angular
ANHYD	Anhydrite
ANTH	Anthracite
APPX	Approximately
ARG	Argillaceous
ARK	Arkose(-ic)
ASHF	Ash flow
ASPH	Asphalt(-ic)
ATTR	Attritus(-al)
AVAIL	Available

AVG	Average
BND	"Band(-s,-ed)"
BS	Base(-al)
BAUX	Bauxite(-itic)
BCM	"Become(-s,-ing)"
BD	"Bed(-ded,-s)"
BDG	Bedding
BLW	Below
BNT	Bentonite
BTW	Between
BIOC	Bioclastic
BITRB	Bioturbated(-ion)
BITN	Bitumen
BIT	Bituminous
BLK	Black
BLKY	Blocky
BLM	Bloom
BN	Bone
BNY	Boney
BTM	Bottom
BLDR	Boulder(-s) (>256mm)
BRCH	Brachiopod
BRAK	Brackish
BREC	Breccia(-ted)
BRT	Bright
BRTL	Brittle
BRKN	Broken
BRN	Brown
BRW	"Burrow(-s,-ed)"
CALC	Calcareous
CLCT	Calcite
CANL	Cannel(-oid)
CARB	Carbonaceous
CAS	Casing
CV	"Caved(-ing,-ings)"
CMT	Cement(-ed)
CM	Centimeter(-s)
CH	Channel
CHRND	Churned
CHRT	Chert(-y)
CLRN	Clarain
CLY	Clay(-ey)(<1/256mm)
CLSH	Clayshale
CLST	Claystone

CLN	Clean
CLT	Cleat(-s)
CB	Cleat butt strike and dip
CF	Cleat face strike and dip
CL	Clinker(-ed)
CLUS	Cluster
CRS	Coarse(-ly)
CSPR	Coal spar(-s)
COALY	Coaly
CBBL	Cobble(64-256mm)
CLUV	Colluvial
COL	Color(-ed)
CLMN	Columnar
COM	Common
CPCT	Compact
CON	Concealed
CONCH	Conchoidal
CNCR	Concretion(-ary)
CIC	Cone-in-cone
CGL	Conglomerate
CNDT	Conodont(-s)
CNSD	Consolidate(-ed)
CTC	Contact
CONTM	Contaminated
CONT	Continuous
CTRT	Contorted
CNVT	Convolute(-ed)
COQ	Coquina
CRL	Coral
CORELST	Corelost(loss)
COV	"Cover(-ed,-ing)"
CR	Cream
CREN	Crenulated
CRND	Crinoid
CTFL	Cutandfill
CTG	Cutting
CRNK	Crinkled
DK	Dark
DEB	Debris
DECR	Decrease(-ing)
DFM	"Deform(-ed,-ation)"
DNS	"Dense(-er,-ly)"
DPT	"Deposit(-ed,-s)"
DPT	"Deposit(-ed,-s)"
DTC	Depth to coal

DTR	Detritus(-al)
DIA	Diameter
DDH	Diamond drillhole
DTM	Diatom
DRTY	Dirty
DIF	"Different(-iated,-ence)"
DCNT	Discontinuous
DISM	Disseminated
DSTD	Distorted
DSTB	Disturbed
DOL	Dolomite(-ic)
DWN	Down(-ward)
DR	Drift
DRLG	Drilling
DH	Drillhole
DURN	Durain
ERTH	Earth(-y)
E	East
ELOG	Electric log
ELV	Elevation
ELIP	Ellipsoid(-al)
ELON	Elongate
EST	Estimate(-ed)
EVAP	Evaporite(-ic)
FNT	Faint(-ly)
FLT	Fault(-ed)
FT	Feet(foot)
FELD	Feldspar(-thic)
FE	Iron-ferruginous
FEST	Ironstone
FLD	Field
FL	Fill(-ed)
F	Fine(-ly)
FCLY	Fireclay
FIS	Fissile
FLKY	Flaky
FLSR	Flaser
FLAT	Flattened
FTCY	Flintclay
FLOR	Fluorescence
FOL	Foliated
FM	Formation
FOSS	Fossil(-iferous)
FRAC	Fracture(-ed)

FRAG	"Fragment(-s,-al,-ed)"
FREC	Frequent
FRSH	Fresh
FRI	Friable
FROS	Frosted
FS	Fusain
FUSL	Fusulinid
GLOG	Gammalog
GSTR	Gastropod
GL	Glass(-y)
GLAUC	Glauconite(-ic)
GRDL	Gradational
GRD	"Grade(-ed,-s)"
GR	"Grain(-ed,-s)"
GRAN	Granule(2-4mm)
GRVL	Gravel(-ly)(~2-76mm)
GY	Gray
GYWKE	Graywacke
GRTR	Grazing trails
GN	Green
GRTY	Gritty
GP	Group
GYP	Gypsum
HKY	Hacky
HD	Hard
HVY	Heavy
HEM	Hematite(-ic)
HI	Highly
HZN	Horizon
DH	Drillhole
HOR	Horizontal
HUM	Humus(-ic)
HC	Hydrocarbon
IGN	Igneous
IMBD	Imbedded
IMPR	Impression
IMP	"Impure(-ity,-ities)"
IN	Inch(-es)
INC	Inclined
INCL	"Includes(-ed,-ing,-sive,-sion"
INCR	"Increase(-ing),-sion)"
IND	Indurated
INBD	Interbedded

INTLAM	Interlaminated
INTSTL	Interstitial
INTV	Interval
INVT	Invertebrate
FE	Iron-ferruginous
FEST	Ironstone
IRREG	Irregular
JNT	Joint(-s)
KAOL	"Kaolin(-ite,-itic)"
LAC	Lacustrine
LAM	Laminated(-ae)
LRG	Large(-er)
LYR	"Layer(-ed,-s)"
LCHD	Leached
LDR	Leader
LEN	Lentil(-icular)
LT	Light(-er)
LIG	Lignite(-itic)
LS	Limestone
LMN	Limonitic
LIN	Lineation
LTHC	Lithic
LTL	Little
LOC	Location
LOES	Loess
LNG	Long
LWR	Lower
LB	Lower bench
LSP	Lower split
LSTR	Lustre
MAGN	Magnetic
MAR	Marine
MRKR	Marker
MRLST	Marlstone
MASS	Massive
MTX	Matrix
MAT	"Matter, material"
MEA	Measured
MED	Medium
MGSP	Megaspores
MBR	Member
MANTH	Meta-anthracite

META	Metamorphic
MICA	Mica(-ceous)
MFLT	Microfault
MICROFOS	Microfossil(-iferous)
MID	Middle
MB	Middle bench
MSP	Middle split
MIDL	Midlustrous
MM	Millimeter(-s)
MIN	"Mineral(-s,-ized)"
MNR	Minor
MIS	Missing
MXD	Mixed
MOD	Moderate
MOT	Mottled
MDFL	Mudflow
MDST	Mudstone
MUSC	Muscovite(-itic)
NR	"No record, not logged, no data, not available"
NOD	"Nodule(-s,-ar)"
N	North
NP	Not present
NTBK	Notebook
NUM	Numerous
OCC	Occasional
OOL	Oolite(-c)
ORNG	Orange
OPQ	Opaque
OVRBRDN	Overburden
OVLY	Overlying
OLV	Olive
OSTR	Ostracod
PALN	Palynology
PT	Part(-ly)
PTG	Parting(-s)
PNL	Parting(s)not located in coalbed
PRLY	Pearly
PBL	Pebble(-s) (4-64mm)
PBLY	Pebbly
PLCY	Pelecypod
PERM	Permeability
PTRW	Petrified wood
PET	Petroleum(-iferous)

PHOS	Phosphate
PNK	Pink
PISO	Pisolite(-ic)
PIT	Pitted
PLAG	Plagioclase
PLAS	Plastic
PL	Plants
PL	Plants
PLTY	Platy
PR	Poor(-ly)
POR	Porous(-sity)
PORPH	Porphyry(-itic)
POSS	Possible(-ility)
PRED	Predominate(-ly)
PRES	Preserved(-ation)
PRIM	Primary
PROB	Probable(-ly)
PROJ	Project
PROM	Prominent(-ly)
PSDO	Pseudo
PRPL	Purple
PYR	"Pyrite(-ic,-ized)"
QTZ	Quartz
QTZTE	Quartzite
QTZSE	Quartzose
QUAT	Quaternary (use QUAT in Comments and Misc Notes Fields only)
RNG	Range(-ing)
RECOV	Recover(-ed)
REG	Regular
RMN	Remains(-nant)
REPL	"Replaced(-ing,-ment)"
RSID	Residual(-uum)
RHY	Rhyolite
RDR	Rider
RPL	Ripple
RPLM	Ripplemark(-s)
RK	Rock
RF	Roof
RFRK	Roofrock
R&P	Room and pillar
ROOT	"Root(-s,-let,-ed)"
RND	Round(-ed)
RUST	"Rust(-ed,-y)"

SACH	Saccharoidal
S&P	Salt and pepper
SMPL	Sample
SD	Sand(1/16-2mm)
SDRK	Sandrock
SS	Sandstone
SDY	Sandy
SCAT	Scattered
SCHST	Schist
STRK	Seatrock
SEC	Secondary
SECT	Section
SED	"Sediment(-s,-ary)"
SEL	Selenite
SANTH	Semianthracite
SEPT	Septarian
SHFT	Shaft
SH	Shale
SHLY	Shaly
SHRP	Sharp
SHL	Shell(-s)
SID	Siderite(-ic)
SIL	Silica(-eous)
SLT	Silt(<1/16mm)
SLST	Siltstone
SLTY	Silty
SKS	Slickensided
SL	Slight(-ly)
SLMP	Slump(-ed)
SML	Small
SMCT	Smectite
SM	Smooth
SPST	Soapstone
SOL	Solution
SRTD	Sorted
SO	South
SPAR	Spar(-s)
SPHL	Sphalerite
SPH	Spherules
SPLNTY	Splintery
SP	Split
STN	"Stain(-ed,-ing)"
STRAT	Strata(-ified)
STR	"Streak(-ed,-s)"
STRI	Striated
STRG	Stringer

STRUC	Structure
STYL	Stylolite
SBANG	Subangular
SUBBIT	Subbituminous
SBHZ	Subhorizontal
S	Sulfur
SURF	Surface
TAB	Tabular
TEX	Texture
THK	Thick(-ness)
THN	Thin
THRU	Throughout
TT	Tight(-ly)
TNG	Tongue
TNST	Tonstein
TCE	Top of Coal Elevation
TGH	Tough
TR	Trace
TRCT	Tract
TRANS	Transitional
TRNSL	Translucent
TRNSP	Transparent
TUB	Tubular
TUF	Tuffaceous
TURQ	Turquoise
UNCONF	Unconformity(-able)
UNCONS	Unconsolidated
UNCSMAT	Unconsolidated material
UC	Underclay
UL	"Underly(-ing,-lain)"
UNDT	Undetermined
UNDIF	Undifferentiated
UNDUL	"Undulates(-d,-ory,-ion)"
UNI	Uniform
UNKN	Unknown
UPR	Upper
UB	Upper bench
USP	Upper split
VAR	"Variable, varies"
VCOL	Varicolored
VRGT	Variegated
VRVD	Varved
VEG	Vegetation(-ed)

VN	Vein
VERT	Vertebrate
VRTL	Vertical
V	Very
VES	Vesicular
VT	Vitrain
VIT	Vitreous
VTRC	Vitric
VTRO	Vitrophyre
VOLC	Volcanic(-s)
WTR	Water
WVY	Wavy
WXY	Waxy
WK	Weak
WTHRD	Weathered
WLD	Welded
WELL	Well
W	West
WHT	White
W-	With
X	Cross (as prefix)
XBD	Cross-bedded
XBDG	Cross-bedding
XLAM	Cross-laminated
XL	"Crystal(-s,-ized)"
XLN	Crystalline
YLW	Yellow
ZN	Zone
&	And
-	Hyphen(not minus sign)
.	Decimal pt(not period-periods should NOT be present)
?	Questionable